

# 予測結果の評価に関する一考察

鈴木 啓 祐

## I は し が き

予測作業をした際に、その作業から得られた予測値がどの程度質のよいものであったか、いいかえれば、どれ程よく将来の値を予測し得たかということを定量的に評価することができれば、その予測作業に用いられた予測方法を定量的に評価することが可能となる。

予測値の評価は、一般に、予測値と実現値とを対比させることによっておこなわれる。これまでに、数多くの評価方法が提唱されて来た。ここでは、そのうち、いくつかの主要なものを挙げ、さらに、この種の評価方法の一つとしてのタイル (H. Theil) の不一致係数 (inequality coefficient) についていくらかの考察をおこなってみた。

## II 各種の予測値の評価方法

2つの変数 (変量) の相関係数の強さは、相関係数 (correlation coefficient)  $r$  で示される。予測値が実現値によく適合しているかどうかは、一見、予測値および実現値という2つの変量の相関関係の強さによって、したがって、相関係数によって、判断されると考えられるが、タイルも指摘するように<sup>1)</sup>、相関係数の高いことが必ずしも、予測値と実測値の近似の程度の高いことを意味しない。

いうまでもなく、変量  $X$  および  $Y$  の相関係数  $r$  は、

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

によって定義される。ただし、 $X_i$  および  $Y_i$  は、それぞれ変量  $X$  および  $Y$  の第  $i$  番目の観察値、 $\bar{X}$  および  $\bar{Y}$  は変量  $X$  および  $Y$  の平均値、 $n$  は観察値の総個数である。

いま、 $X_i$  と  $Y_i$  ( $i=1, 2, \dots, n$ ) との間に、

$$Y_i = X_i \quad (2.2)$$

という関係が成立しているとすれば、

$$\bar{X} = \bar{Y} \quad (2.3)$$

であるから、

$$\begin{aligned} r &= \frac{\sum (X_i - \bar{X})(X_i - \bar{X})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (X_i - \bar{X})^2}} \\ &= 1 \end{aligned} \quad (2.4)$$

となる。

また、もし、 $X_i$  と  $Y_i$  との間に、

$$Y_i = a + bX_i \quad (2.5)$$

という関係があれば、

$$\bar{Y} = a + b\bar{X} \quad (2.6)$$

であるから、

$$\begin{aligned} r &= \frac{\sum (X_i - \bar{X})\{(a + bX_i) - (a + b\bar{X})\}}{\sqrt{\sum (X_i - \bar{X})^2 \sum \{(a + bX_i) - (a + b\bar{X})\}^2}} \\ &= 1 \end{aligned} \quad (2.7)$$

となる。

すなわち、相関係数は  $Y_i = X_i$  という関係がなく、 $Y_i = a + bX_i$  という関係が存在する場合でも1となることが可能である。

したがって、上記の  $X_i$  を予測値、 $Y_i$  を実測値とするとき、予測値とその実測値とが一致をしない場合でも、それらの間に直線的関係がありさえすれば、それらの間の相関係数は1となるといえる。それゆえ、相関係数は、予測値と

1) H. Theil, *Economic Forecasts and Policy*, Amsterdam, North-Holland Publishing Company, 1961, p. 31. 岡本哲治訳『アンリ・タイル著 経済の予測と政策』東京、創文社、昭和39年、40頁。

実現値との間の近似性を示すよい指標であるとはいえない。

予測値と実現値との間の近似性の判定の指標の1つとして、まず、percentage of verification  $P$  が挙げられる<sup>2)</sup>。

この指標は、気象学の分野で用いられているもので、

$$P = \frac{\text{全予測成功数}}{\text{全予測回数}} \times 100 \quad (2.8)$$

という式で定義され得る値である。これをより具体的に示せば、つぎのようになる。いま、起り得る気象の種類を  $r$  個とし、予測結果が第  $i$  番目の気象であり、実際に起った気象が第  $j$  番目のものであった回数を  $n_{ij}$  としたとき、 $P$  は、

$$P = \frac{\sum_{i=1}^r n_{ii}}{n_{..}} \times 100 \quad (2.9)$$

によって示される。ただし、 $n_{..}$  は全予測回数、あるいは  $\sum_{j=1}^r \sum_{i=1}^r n_{ij}$  である。

$n_{ij}$  を1つの表によって明示してみると、第1表のようになる。

第1表  $n_{ij}$  の一覧表

実測結果 予測結果	1	2.....j.....r	合 計
1	$n_{11}$	$n_{12} \dots n_{1j} \dots n_{1r}$	$n_{1.}$
2	$n_{21}$	$n_{22} \dots n_{2j} \dots n_{2r}$	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$i$	$n_{i1}$	$n_{i2} \dots n_{ij} \dots n_{ir}$	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$r$	$n_{r1}$	$n_{r2} \dots n_{rj} \dots n_{rr}$	$n_{r.}$
合 計	$n_{.1}$	$n_{.2} \dots n_{.j} \dots n_{.r}$	$n_{..}$

(注)  $n_{i.} = \sum_{j=1}^r n_{ij}$ ,  $n_{.j} = \sum_{i=1}^r n_{ij}$ ,  $n_{..} = \sum_{j=1}^r \sum_{i=1}^r n_{ij}$

ヴァレン (A. Wallén) は、1921年、連関係数  $V$  が予測のよさの尺度として用いられ得ることを提唱した。ヴァレンのいう  $V$  とは、

$$V = \frac{n_{11}n_{22} - n_{12}n_{21}}{\sqrt{n_{.1}n_{.2}n_{2.}n_{1.}}} \quad (2.10)$$

2) 鈴木栄一「天気予報に対する Verification について

(1)『予報研究ノート』第3巻第9号、1953年、9-23頁。

で定義される値である<sup>3)</sup>。ただし、 $n_{ij}$  ( $i=1,2$ ;  $j=1,2$ ) は、第1表における  $n_{ij}$  と同じ意味をもつ記号であるが、 $V$  の算出においては、 $r$  は2とされている。

鈴木栄一氏は、一般に、

$$\chi^2 = \sum_{j=1}^s \sum_{i=1}^r \frac{\left(n_{ij} - \frac{n_{i.}n_{.j}}{n_{..}}\right)^2}{\frac{n_{i.}n_{.j}}{n_{..}}} \quad (2.11)$$

が自由度  $(r-1)(s-1)$  の  $\chi^2$  分布に従うことから、この値  $\chi^2$  あるいは  $\chi$  を予測値の良否の指標とし得ることを明らかにした<sup>4)</sup>。ただし、 $r$  とは気象の予測結果の全種類、 $s$  とは実現された結果の全種類である。したがって、この  $\chi^2$  は、 $r$  種の予測結果と、 $s$  種 ( $s \neq r$ ) の実現された結果とが存在する場合に利用し得る指標であり、きわめて利用範囲の大きな指標であるといえよう。

これに対して、交通予測の分野においても予測値と実現値との間の差を測定する指標が考案されている。その値は、RMS 誤差 (RMS error) と呼ばれるものである<sup>5)</sup>。この値は、

$$\text{RMS 誤差} = \sqrt{\frac{\sum_{i=1}^n (\hat{Q}_i - Q_i)^2}{n}} \quad (2.12)$$

によって定義される。ただし、この値は、いくつかの経路があり、その各経路の交通量を予測することを前提として定義された指標であり、 $\hat{Q}_i$  は、第  $i$  経路の交通量の予測値、 $Q_i$  はその実現値、 $n$  は交通量の予測の対象となった経路の全個数である。この指標は、気象の予測結果 (定量的予測結果) の評価の指標としても用いら

3) 鈴木栄一「天気予報に対する Verification について (2)」『予報研究ノート』第4巻第1号、1953年、1-25頁。

Axel Wallén, "Sur la Cantrole des Annonces de Tempêtes," "Geografiska Annaler, 3, 1921, pp. 267-277.

4) 鈴木栄一、前掲論文、(第4巻第1号)20頁。

5) Arther B. Soslaw and Glenn E. Brokke, "Appraisal of Sample Size based on Phoenix O-D Survey Data," Highway Research Board, Bulletin, No. 253, Washington, Highway Research Board, 1960, pp. 114-127.

広島都市交通問題研究会編『広島都市交通の現況と将来』東京、大蔵省印刷局、1971年、219-220頁。

れている<sup>6)</sup>。

このときは、 $\hat{Q}_i$  と  $Q_i$  は、第1表とはやや異なった第2表のような形で整理される。すなわち、この表の中には、出現度数ではなく、予測結果そのものが記入される。

第2表  $\hat{Q}_i$  および  $Q_i$  の一覧表

予測値番号	予 測 値	実 現 値
1	$\hat{Q}_1$	$Q_1$
2	$\hat{Q}_2$	$Q_2$
...	...	...
$i$	$\hat{Q}_i$	$Q_i$
...	...	...
$n$	$\hat{Q}_n$	$Q_n$

形式的には、上記の式 (2.12) とよく類似しているが、その質がまったく異なったつぎのような指標  $C$  をも予測値評価の指標として用いることができる。すなわち、

$$C = \sum_{i=1}^n \frac{(\hat{F}_i - F_i)^2}{F_i} \quad (2.13)$$

で定義される指標  $C$  は、一種の予測値評価の指標として用いられ得る。ただし、 $\hat{F}_i$  とは、予測値の第  $i$  番目の階級の得られた回数、 $F_i$  とはその階級の値をもつ実現値の得られた回数である。これは、交通量の予測値の評価に用いられている<sup>7)</sup>。このときの  $\hat{F}_i$  と  $F_i$  は、第3表のような表で示される。

第3表  $\hat{F}_i$  および  $F_i$  の一覧表

階 級	度 数	
	予 測 値	実 現 値
$C_1$	$\hat{F}_1$	$F_1$
$C_2$	$\hat{F}_2$	$F_2$
...	...	...
$C_i$	$\hat{F}_i$	$F_i$
...	...	...
$C_n$	$\hat{F}_n$	$F_n$
合 計	$\Sigma \hat{F}_i$	$\Sigma F_i$

(注)  $C_i$  とは 第  $i$  階級の階級代表値である。

この指標  $C$  は必ずしもよい指標であるとはい

えない。その理由は、(1)予測値と実現値との完全な一致が見られなくとも、それらの差異が十分小さければ、予測値の度数分布と実現値のそれとは互いに同一のものとなり得るし、また、(2)予測値が実現値とまったく異なっているものの度数分布が偶然一致することがあるからである。しかし、(1)のような状態が起っている場合、度数分布が同一となる程度の不一致は不一致とみなさないならば、この指標は、実用的には、予測値評価のよい指標とみなし得るであろう。しかも、この指標は、自由度  $(n-1)$  の  $\chi^2$  分布に従うから、この性質を用いれば、この指標によって予測値に対する統計的仮説検定をおこなうことができる。この点から見て、この指標は、式 (2.11) の  $\chi^2$  と同様に統計的仮説検定の可能な指標といえる。

計量経済学の分野では、タイルによって、彼の不一致係数が提唱された。タイルの不一致係数  $U$  は、

$$U = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - A_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n P_i^2} + \sqrt{\frac{1}{n} \sum_{i=1}^n A_i^2}} \quad (2.14)$$

によって定義される  $U$  である。ただし、 $P_i$  は第  $i$  番目の予測値 ( $P$ )、 $A_i$  は第  $i$  番目の実測値 ( $A$ )、 $n$  は予測値の全個数である<sup>8)</sup>。この  $U$  は 0 から 1 までの値をとり、0 に近い程、予測値と実現値の差が小さいことを示す。

タイルの不一致係数の持つ重要な性質は、不一致係数の自乗  $U^2$  が、 $U_M^2$ 、 $U_S^2$ 、および  $U_C^2$  の3つの部分に分けられること、すなわち、

$$U^2 = U_M^2 + U_S^2 + U_C^2 \quad (2.15)$$

となることである。ただし、

$$U_M = \frac{\bar{P} - \bar{A}}{D}, \quad U_S = \frac{S_P - S_A}{D},$$

6) Glenn Brier and Roger A. Allen, "Verification of Weather Forecasts," in Thomas F. Malone ed., *Compendium of Meteorology*, Boston, American Meteorological Society 1951, pp. 841-848.

鈴木栄一, 前掲論文 (第4巻, 第1号)。

7) 広島都市交通問題研究会編, 前掲書, 219-220頁。

8) H. Theil, *op. cit.*, pp. 31-42. 岡本哲治訳, 前掲書, 40-52頁。

$$U_C = \frac{\sqrt{2(1-r)}S_P S_A}{D}$$

である。これらの式で、 $\bar{P}$  および  $\bar{A}$  は、それぞれ  $P$  および  $A$  の平均値 (算術平均)、 $S_P$  および  $S_A$  は、 $P$  および  $A$  の標準偏差、 $r$  は  $P$  および  $A$  の相関係数、そして、 $D$  は  $U$  の分母である。したがって、

$$U^M = \frac{U_M^2}{U^2} \quad (2.16)$$

$$U^S = \frac{U_S^2}{U^2} \quad (2.17)$$

$$U^C = \frac{U_C^2}{U^2} \quad (2.18)$$

とすれば、

$$U^M + U^S + U^C = 1 \quad (2.19)$$

となる。

タイルは、 $U^M$  を偏りの構成割合 (bias proportion)、 $U^S$  を分散の構成割合 (variance proportion)、そして、 $U^C$  を共分散の構成割合 (covariance proportion) と名づけた。そして、これらを3つの異なった原因による不等度の構成割合 (proportions of inequality) と呼んだ。その理由は、彼によれば、 $U_M$  が中心値 (すなわち、平均値) の不等 (unequal central tendency) に基づく偏不等度 (partial coefficient of inequality)、 $U_S$  が変動の不等 (unequal variation) に基づく偏不等度、 $U_C$  が共変関係の不完全 (imperfect covariation) に基づく偏不等度とみなし得るからである<sup>9)</sup>。

さらにタイルは、 $U^2$  を、

$$U^2 = U_M^2 + U_R^2 + U_D^2 \quad (2.20)$$

のように示すことのできることを明らかにした。ただし、

$$U_R = \frac{S_P - rS_A}{D}, \quad U_D = \frac{\sqrt{(1-r^2)}S_A}{D}$$

である。そして、 $U_R$  を回帰係数不良 (incorrect regression slope) による不等度、 $U_D$  を回帰の乱れ (nonzero regression disturbance) による不等度と名づけた。また、

9) H. Theil, *op. cit.*, pp. 31-38. 岡本哲治訳, 前掲書, 40-47頁。

$$U^R = \frac{U_R^2}{U^2} \quad (2.21)$$

$$U^D = \frac{U_D^2}{U^2} \quad (2.22)$$

とすれば、

$$U^M + U^R + U^D = 1 \quad (2.23)$$

という関係が成立する。彼は、 $U^R$  を回帰係数の構成割合 regression proportion、 $U^D$  を乱れの構成割合 disturbance proportion と名づけた<sup>10)</sup>。

このようなタイルの不一致係数の他に、情報理論を用いた予測値の評価のための指標が工夫されている。この指標は、

$$I = \sum_{i=1}^n a_i \log \frac{a_i}{p_i} \\ = \sum_{i=1}^n a_i \log \frac{1}{p_i} - \sum_{i=1}^n a_i \log \frac{1}{a_i} \quad (2.24)$$

で定義される  $I$  である。ただし、 $\sum_{i=1}^n p_i = 1$ ,  $\sum_{i=1}^n a_i = 1$  であり、 $p_i$  ( $i=1, 2, \dots, n$ ) は第  $i$  番目の事物の prior probability,  $a_i$  ( $i=1, 2, \dots, n$ ) は第  $i$  番目の事物の観察値から得られた出現確率である。

この指標を紹介したブレナン (M. J. Brennan) とタフト (Philip Taft), およびシュパック (Mark B. Schupack)<sup>11)</sup> は期待された結果のもつ情報量が  $a_i$  を得ることによってどれだけ減じたかを示す値 (reduction in expected information caused by receiving the particular message in the  $a_i$ 's) であると述べている<sup>12)</sup>。

これらが、予測値の評価の指標の簡単な概観であるが、これらの指標は、2種の指標に大別され得る。その1つは、「予測の成功度に関する指標」とでもいえる指標、すなわち、予測が成功したか否かという判定の存在を前提として予測値の評価をおこなう指標、他の1つは、「予測値そのものの良さの指標」である。前者

10) H. Theil, *op. cit.*, pp. 38-42. 岡本哲治訳, 前掲書, 47-52頁。

11) M. J. Brennan, Philip Taft, and Mark B. Schupack, *The Economics of Age*, New York, W. W. Norton, 1967, pp. 182-188.

12) M. J. Brennan, Philip Taft, and Mark B. Schupack, *op. cit.*, p. 185.

には、式(2.8)あるいは(2.9)で示される $P$ 、式(2.10)の $V$ 、および式(2.11)の $\chi^2$ があり、後者には、RMS誤差、式(2.13)の指標 $C$ 、タイルの不一致係数、および指標 $I$ である。

### III 予測誤差と予測誤差率

予測値と実現値との不一致の程度を定量的に表現する方法には2つの方法がある。まず第1の方法はこれらの値の差である。一般に、第 $i$ 番目の予測値 $P_i$ とその実現値 $A_i$ との不一致を示すための差 $d_i$ は、

$$d_i = P_i - A_i \quad (3.1)$$

あるいは、

$$d_i = A_i - P_i \quad (3.2)$$

さらに、また、

$$d_i = |A_i - P_i| \quad (3.3)$$

で示されるであろう。式(3.1)の $d_i$ は実測値に対する予測値の差異であり、式(3.2)の $d_i$ は、予測値に対する実測値の差異であると考えられる。そして、式(3.3)の $d_i$ は予測値と実測値との不一致の大きさを示す値であるといえる。

これに対して、比率によっても、それらの不一致を示すことができる。すなわち、第 $i$ 番目の予測値とその実測値との不一致を示すための比率 $p_i$ は、

$$p_i = \frac{P_i}{A_i} \quad (3.4)$$

または、

$$p_i = \frac{A_i}{P_i} \quad (3.5)$$

によって示されるであろう。式(3.4)は、実測値に対する予測値の相対的差異、式(3.5)は予測値に対する実測値の相対的差異であるといえる。

ここでは、 $d_i$ を予測誤差、 $p_i$ を予測誤差率と呼ぶことにする<sup>13)</sup>。そして、 $d_i$ としては、式(3.1)を、また $p_i$ としては式(3.4)を用いることにする。

なお、 $p_i$ をその対数によって示せば、式(3.

4) および(3.5)は、

$$\log p_i = \log P_i - \log A_i \quad (3.6)$$

および、

$$\log p_i = \log A_i - \log P_i \quad (3.7)$$

となる。そして、このとき、 $\log p_i$ の絶対値は、

$$|\log p_i| = |\log P_i - \log A_i| \quad (3.8)$$

であるから、これは、予測値と実現値の各対数の不一致の大きさを示す値であるといえる。

また、予測誤差率によって予測値と実現値の不一致を示すことのできる場合は、つねに、 $A_i > 0$ 、 $P_i > 0$ となり得るときのみである。

このように、予測値と実現値の不一致を示す値を分類してみると、タイルの不一致係数は、ここでいう予測誤差に関する指標であるといえる。

### IV タイルの不一致係数の拡張

予測誤差 $d_i$ は、

$$d_i = P_i - A_i \quad (4.1)$$

で示される。したがって、このとき、

$$P_i = A_i + d_i \quad (4.2)$$

となる。しかし、 $d_i$ は必ずしも0のまわりに分布しないので、 $P_i$ と $A_i$ との関係を示す直線は必ずしも、

$$P_i = A_i \quad (4.3)$$

のようにならない。タイルは、 $P_i$ と $A_i$ との関係は、一般に、

$$P_i = a + bA_i + \varepsilon_i \quad (4.4)$$

によって示されると考えた<sup>14)</sup>。ただし、 $a$ および $b$ はパラメーター、 $\varepsilon_i$ は残差である。この式では、

$$d_i = a + (b-1)A_i + \varepsilon_i \quad (4.5)$$

とみなされていると考えられる。なんとなれば、この $d_i$ を式(4.2)に代入すれば、

$$P_i = A_i + a + (b-1)A_i + \varepsilon_i$$

14) H. Theil, *op. cit.*, pp. 31-32. 岡本哲治, 前掲書, 40頁.

なお、タイルは、実際に、

$$P_i = a + bA_i + \varepsilon_i \quad (a)$$

という式を示していず、

$$P_i = a + bA_i \quad (b)$$

の成立することに関して論じているが、彼の議論では、明らかに、式(a)のような関係を前提としている。

13) 誤差の1つとして標本誤差、すなわち、標本と母集団との間の誤差がある。この誤差と予測値と実現値との間の誤差とを区別するために、後者に「予測誤差」という名称を与えた。

$$=a+bA_i+\varepsilon_i \quad (4.6)$$

となり，式(4.4)が得られるからである。

他方，予測誤差率  $p_i$  の対数は，

$$\log p_i = \log P_i - \log A_i \quad (4.7)$$

であるから，

$$\log P_i = \log A_i + \log p_i \quad (4.8)$$

となる。このとき，

$$\log p_i = a + (b-1) \log A_i + \log e_i \quad (4.9)$$

と仮定してみよう。ただし， $a$  および  $b$  はパラメーター， $e_i$  は残差である。この仮定は，式(4.5)の  $d_i$  を  $\log p_i$ ， $A_i$  を  $\log A_i$  として  $\varepsilon_i$  を  $\log e_i$  で置きかえて得られたものである。この仮定を用いると，

$$\log P_i = a + b \log A_i + \log e_i \quad (4.10)$$

が得られる。この式は，

$$P_i = k A_i^b e_i \quad (4.11)$$

を対数で示した式に他ならない。ただし， $k = \log a$  である。

いま， $P_i$  あるいは  $\log P_i$  を  $\pi_i$  で， $A_i$  あるいは  $\log A_i$  を  $\alpha_i$  で，また， $\varepsilon_i$  あるいは  $\log e_i$  を  $\delta_i$  で示せば，式(4.4)と式(4.10)とは，

$$\pi_i = a + b\alpha_i + \delta_i \quad (4.12)$$

という1つの式で示される。

この式(4.12)は， $P_i$  のモデルとして式(4.4)を採用するとき， $\pi_i = P_i$ ， $\alpha_i = A_i$ ， $\delta_i = \varepsilon_i$  となり，式(4.11)を採用するとき， $\pi_i = \log P_i$ ， $\alpha_i = \log A_i$ ， $\delta_i = \log e_i$  となるといえる。

いま，タイルの不一致係数(式(2.14))において， $P_i$  を  $\pi_i$ ， $A_i$  を  $\alpha_i$  で置きかえた係数を  $U^*$  とすれば， $U^*$  は，

$$U^* = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (\pi_i - \alpha_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n \pi_i^2} + \sqrt{\frac{1}{n} \sum_{i=1}^n \alpha_i^2}} \quad (4.13)$$

となる。 $\pi_i$  および  $\alpha_i$  の定義によれば，この  $U^*$  から2種類の  $U^*$ ，すなわち， $U$  および  $U'$  を得ることができる。すなわち，

$$(1) \quad U = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - A_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n P_i^2} + \sqrt{\frac{1}{n} \sum_{i=1}^n A_i^2}} \quad (4.14)$$

$$(2) \quad U' = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (\log P_i - \log A_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (\log P_i)^2} + \sqrt{\frac{1}{n} \sum_{i=1}^n (\log A_i)^2}} \quad (4.15)$$

$$= \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n \left(\log \frac{P_i}{A_i}\right)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (\log P_i)^2} + \sqrt{\frac{1}{n} \sum_{i=1}^n (\log A_i)^2}} \quad (4.15')$$

である。

式(4.14)で示される  $U$  はタイルの不一致係数に他ならない。これは， $U'$  に対して，予測誤差の不一致係数(coefficient of inequality in difference)といえよう。また，式(4.15)あるいは(4.15')で示される  $U'$  は予測誤差率の不一致係数(coefficient of inequality in proportion)というべきものであろう。

この係数  $U^*$  は，タイルの不一致係数と同一の性質が見られる。すなわち，まず，

$$0 \leq U^* \leq 1 \quad (4.16)$$

である。

さらに，係数  $U^*$  は  $U_M^*$ ， $U_S^*$  および  $U_C^*$  の3つの部分に分けることができる。すなわち，

$$U^{*2} = U_M^{*2} + U_S^{*2} + U_C^{*2} \quad (4.17)$$

と書ける。

ただし，

$$U_M^* = \frac{\bar{\pi} - \bar{\alpha}}{D^*}, \quad U_S^* = \frac{S_{\pi^*} - S_{\alpha^*}}{D^*},$$

$$U_C^* = \sqrt{\frac{2(1-r^*)S_{\pi^*}S_{\alpha^*}}{D^*}}$$

である。これらの式で  $\bar{\pi}$  および  $\bar{\alpha}$  は，それぞれ  $\pi$  および  $\alpha$  の平均， $S_{\pi^*}$  および  $S_{\alpha^*}$  は， $\pi$  および  $\alpha$  の標準偏差， $r^*$  は  $\pi$  と  $\alpha$  との相関係数，そして  $D^*$  は  $U^*$  の分母である。したがって，

$$U_M^{*2} = \frac{U_M^{*2}}{U^{*2}} \quad (4.18)$$

$$U_S^{*2} = \frac{U_S^{*2}}{U^{*2}} \quad (4.19)$$

$$U_C^{*2} = \frac{U_C^{*2}}{U^{*2}} \quad (4.20)$$

とすれば，

$$U_M^{*2} + U_S^{*2} + U_C^{*2} = 1 \quad (4.21)$$

となる。

また,

$$U^{*2} = U_M^{*2} + U_R^{*2} + U_D^{*2} \quad (4.22)$$

という3つの要素の合成物として示すこともできる。ただし,

$$U_R^* = \frac{S_\pi^* - r^* S_\alpha^*}{D^*},$$

$$U_D^* = \frac{(1 - r^{*2})}{D^*} S_\alpha^*$$

である。そして,

$$U_R^{*2} = \frac{U_R^{*2}}{U^{*2}} \quad (4.23)$$

$$U_D^{*2} = \frac{U_D^{*2}}{U^{*2}} \quad (4.24)$$

とすれば,

$$U_M^* + U_R^* + U_D^* = 1 \quad (4.25)$$

という関係が成立する。

## V 予測値の実現値のまわりの分布

$U^*$  によると予測誤差および予測誤差率を定量的に評価することができる。ところで,  $U$  の分布が不明であると同様に  $U^*$  の分布も不明である。しかし, いま,

$$s^{(+)*} = \sqrt{\frac{1}{n^{(+)}} \sum_{i=1}^{n^{(+)}} \delta_i^{(+)*2}} \quad (5.1)$$

$$s^{(-)*} = \sqrt{\frac{1}{n^{(-)}} \sum_{i=1}^{n^{(-)}} \delta_i^{(-)*2}} \quad (5.2)$$

で定義される  $s^{(+)*}$  および  $s^{(-)*}$  をそれぞれ「実現値より大きな予測値の標準偏差」および「実現値より小さな予測値の標準偏差」とすれば, 予測値の実現値のまわりの分布の状態をこれら  $s^{(+)*}$  および  $s^{(-)*}$  によって示すことができる。

ただし,

$$\delta_i^{(+)} = \pi_i - \alpha_i (\geq 0)$$

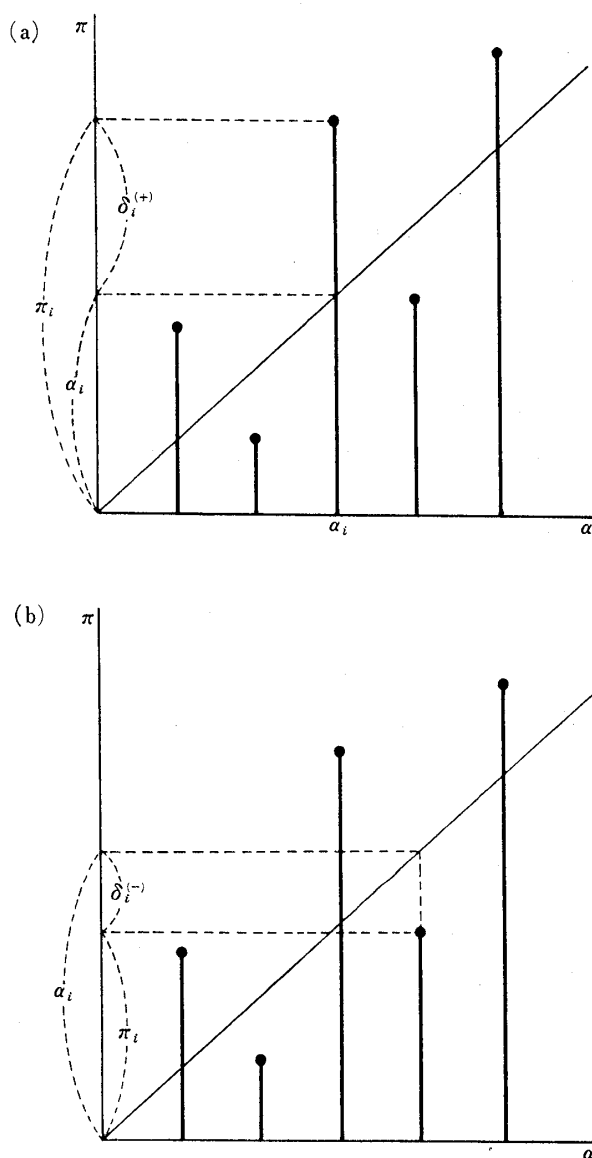
$$\delta_i^{(-)} = \pi_i - \alpha_i (\leq 0)$$

$n^{(+)}$  は  $\delta_i^{(+)}$  の全個数,  $n^{(-)}$  は  $\delta_i^{(-)}$  の全個数である (第1図参照)。

$s^{(+)*}$  は「正の標準偏差 (positive standard deviation)」,  $s^{(-)*}$  は「負の標準偏差 (negative standard deviation)」と名づけられるべき値である。

$s^{(+)*}$  と  $s^{(-)*}$  に対しては, チェビシェフの

第1図  $\delta_i^{(+)}$  および  $\delta_i^{(-)}$  を示す図



不等式 (Tschebychev's inequality), あるいはビアンネメ・チェビシェフの不等式 (inégalité de Bienaymé-Tschebychev)<sup>15)</sup> と同様の不等式が適用され得る。すなわち,

$$\frac{n\lambda^{(+)}}{n^{(+)}} \leq \frac{1}{\lambda^2} \quad (\lambda \geq 1) \quad (5.3)$$

$$\frac{n\lambda^{(-)}}{n^{(-)}} \leq \frac{1}{\lambda^2} \quad (\lambda \geq 1) \quad (5.4)$$

15) 中山伊知郎編『統計学辞典 増補版』東京, 東洋経済新報社, 昭和39年, 80-81頁。

Maurice G. Kendall and Alan Stuart, *The Advanced Theory of Statistics*, Vol. 2, New York, Hafner, 1967, pp. 3-4.

という不等式が成立する<sup>16)</sup>。ただし、 $n\lambda^{(+)}$  とは  $\lambda s^{(+)*}$  より大きな値をもった  $\delta_i^{(+)}$  の個数であり、 $n\lambda^{(-)}$  とは  $\lambda s^{(-)*}$  より小さな値をもった  $\delta_i^{(-)}$  の個数である。

このような性質を用いることによって、予測値の実測値からの具体的な大きさを示すことができる。すなわち、 $\alpha_i - \lambda s^{(-)*}$  から  $\alpha_i + \lambda s^{(+)*}$  までの範囲には全予測値の  $1 - (1/\lambda^2)$  以上の予測値が存在し、たとえば、 $\lambda=3$  としたときには、その範囲の中には、全体の88.9%以上の予測値が存在することになる。これは、いうまでもなく、予測値の分布に無関係に成立する性質であるので、どのような予測結果に対しても一般的に用いられる性質である。

## VI 1つの実例

ここで、1つの実例を挙げて、 $U^*$  の適用結果がどのようなものであるかを示してみよう<sup>17)</sup>。

ここで用いられる実例は、わが国における1962年の地域間物資輸送量の予測値 ( $P_i$ )——これは1960年の資料を用いたものである——の評価である<sup>18)</sup>。

この予測値は、アイザードの引力モデル<sup>19)</sup>を少し変形したモデル<sup>20)</sup>を用いて得られたもので

16) 式(5.3)および(5.4)の成立することは、つぎのようにして証明できる。

まず、定義により、

$$s^{(+)*2} = \frac{1}{n^{(+)} \sum_{i=1}^n \delta_i^{(+)*2}}$$

である。したがって、

$$\begin{aligned} n^{(+)} s^{(+)*2} &= \sum_{i=1}^n \delta_i^{(+)*2} \\ &\geq \sum_{i=1}^n \lambda \delta_i^{(+)*2} \geq n \lambda^{(+)} \lambda^2 s^{(+)*2} \end{aligned}$$

ただし、 $\lambda \delta_i^{(+)*}$  とは  $\delta_i^{(+)*} \geq \lambda s^{(+)*}$  となるような  $\delta_i^{(+)*}$  である。それゆえ、

$$\frac{n\lambda^{(+)}}{n^{(+)}} \leq \frac{1}{\lambda^2}$$

が得られる。同様に考えて、

$$\frac{n\lambda^{(-)}}{n^{(-)}} \leq \frac{1}{\lambda^2}$$

という関係が得られる。

17) ここでの計算は、すべて、青山学院大学電子計算機センターにおいておこなわれた。

18) 鈴木啓祐『物資輸送量の計測と予測』東京、交通日本社、昭和44年、111頁、233頁。

ある。図によってその予測値とその実現値との関係を示してみると、第2図のようになる。この図では  $\pi_i = P_i$ 、 $\alpha_i = A_i$  として画かれたものであると考えられる。いま、 $A_i$  と  $P_i$  との関係を表わしたような1次式で示すと、

$$P_i = -6.6439 + 0.7458 A_i \quad (6.1)$$

( $r=0.8210$ )

となった。ただし、 $r$  は  $A_i$  と  $P_i$  の相関係数である。そして、

$$U = 0.2714 \quad (6.2)$$

$$\left. \begin{aligned} U^M &= 0.129 \\ U^S &= 0.022 \\ U^C &= 0.849 \end{aligned} \right\} (6.3) \quad \left. \begin{aligned} U^M &= 0.129 \\ U^R &= 0.020 \\ U^D &= 0.861 \end{aligned} \right\} (6.4)$$

となった。さらに、

$$\left. \begin{aligned} s^{(+)} &= 69.8 ; & (n^{(+)} &= 16) \\ s^{(-)} &= 127.0 ; & (n^{(-)} &= 40) \\ n &= 56 \end{aligned} \right\} \quad (6.5)$$

となった。ただし、 $s^{(+)}$  および  $s^{(-)}$  とは、 $s^{(+)*}$  および  $s^{(-)*}$  において、 $\pi_i$  を  $P_i$ 、 $\alpha_i$  を  $A_i$  としたときの値である。また、 $n$  は、 $s^{(+)}$  と  $s^{(-)}$  との合計である ( $P_i = A_i$  のような  $P_i$  が存在しないときは  $s^{(+)} + s^{(-)}$  は  $n$  に等しくなるが、 $P_i = A_i$  のような  $P_i$  の存在する場合は、その値は予測値の全個数  $n$  よりも大きくなる)。

ところで、輸送量はつねに0以上の値であるから、理論的に  $\alpha_i - \lambda s^{(-)}$  を算出したとき0以下となるという矛盾が起る。実際、第2図には、

$$P_i = A_i - 3s^{(-)} \quad (6.6)$$

および、

$$P_i = A_i + 3s^{(+)} \quad (6.7)$$

が画かれているが、 $A_i$  が正の部分においても式(6.6)の図は負の部分を通っている。しかし、つぎに述べる  $U'$  においては、このような矛盾は起らない。

第3図は、 $\log A_i$  と  $\log P_i$  との関係を示した図である。このとき、これらの値の関係は、

$$\log P_i = 0.2308 + 0.7883 \log A_i \quad (6.8)$$

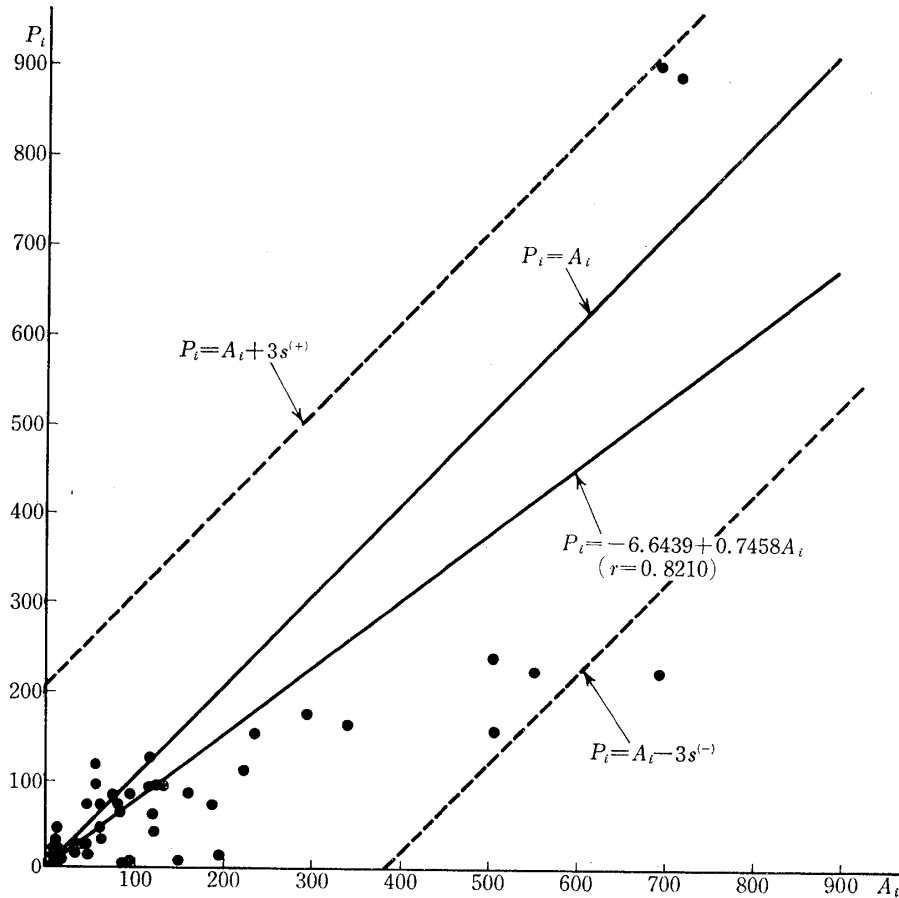
( $r=0.8012$ )

19) Walter Isard, *Methods of Regional Analysis*, Cambridge, Mass, The M.I.T. Press, 1960, pp. 494-497.

20) 鈴木啓祐、前掲書、233頁。



第2図 予測値と実現値との関係



で示される。ただし、 $r$  はここでも相関係数である。そして、

$$U' = 0.1180 \quad (6.9)$$

$$\left. \begin{array}{l} U^{M'} = 0.104 \\ U^{S'} = 0.0006 \\ U^{C'} = 0.895 \end{array} \right\} (6.10) \quad \left. \begin{array}{l} U^{M'} = 0.104 \\ U^{R'} = 0.078 \\ U^{D'} = 0.818 \end{array} \right\} (6.11)$$

である。ただし、 $U^{M'}$ 、 $U^{S'}$ 、 $U^{C'}$  および  $U^{D'}$  は  $U^{M*}$ 、 $U^{S*}$ 、 $U^{C*}$  および  $U^{D*}$  において、 $\pi_i$  を  $\log P_i$ 、 $\alpha_i$  を  $\log A_i$  で示したものである。さらに、

$$\left. \begin{array}{l} s^{(+)' } = 0.356 \\ s^{(-)' } = 0.445 \end{array} \right\} \quad (6.12)$$

である。ただし、 $S^{(+)'}$  および  $S^{(-)'}$  は  $S^{(+)*}$  および  $S^{(-)*}$  において  $\pi_i$  を  $\log P_i$ 、 $\alpha_i$  を  $\log A_i$  とした値である。第3図には、

$$\log P_i = \log A_i - 3s^{(-)' } \quad (6.13)$$

および、

$$\log P_i = \log A_i + 3s^{(+)' } \quad (6.14)$$

を記入しておいたが、このときには、式(6.6)

に関して発生したような矛盾は現われない。

## VII 結 語

これまでに、いくつかの予測値の評価の方法が提唱されて来た。ここでは、そのうち、特に、タイルの不一致係数の拡張を試みた。すなわち、予測誤差と予測誤差率の定量的評価のための指標として、係数  $U^*$  を定義した。

この係数  $U^*$  はその定義に用いられた式の  $\pi_i$  および  $\alpha_i$  を  $P_i$  および  $A_i$  にするか、あるいは、 $\log P_i$  および  $\log A_i$  にするかによって予測誤差の指標  $U$ ——これは、タイルの不一致係数である——または、予測誤差率の指標  $U'$  になり得る。

もし、 $A_i$  と  $P_i$  との間に、

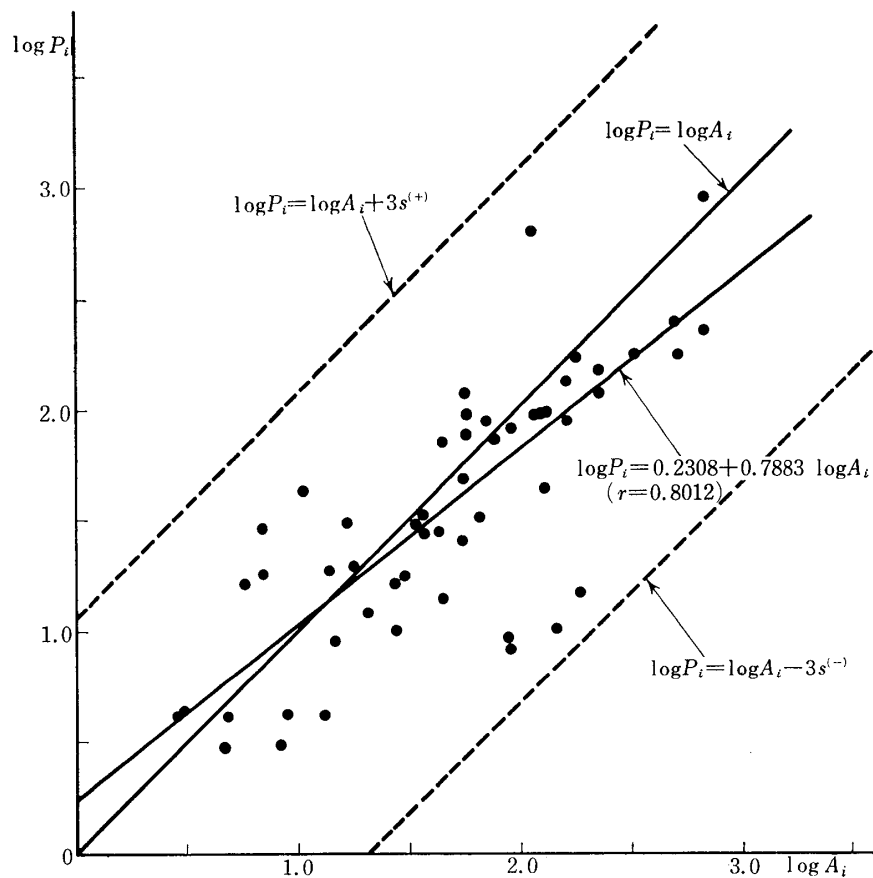
$$P_i = a + bA_i + \varepsilon_i \quad (7.1)$$

が成立し、したがって、

$$P_i \geq 0 \quad (7.2)$$

となるような場合は、 $U$  を用いることが望まし

第3図 予測値の対数と実現値の対数との関係



く、もし、

$$\log P_i = a + b \log A_i + \varepsilon_i \quad (7.3)$$

あるいは、

$$P_i = (kA^b)e_i \quad (7.4)$$

とみなされる関係が成立し、

$$P_i > 0 \quad (7.5)$$

となるような場合は、 $U'$  を用いることが望ましいといえよう。

しかし、たとえ、 $P_i \geq 0$  が成立した場合でも、また、 $P_i > 0$  が成立した場合でも、単に、予測誤差、あるいは、予測誤差率の大きさを測定するという目的のためには、 $U$  および  $U'$  のうちのいずれをも用いることが許されるであろう。

さらに、 $s^{(+)*}$  および  $s^{(-)*}$  を定義することによって、予測値の分布状態を予測値の分布の形に無関係に記述することを試みた。

この  $s^{(+)*}$  および  $s^{(-)*}$  を用いて、予測値分布の範囲を表現する場合、予測値の範囲はかな

り大きく表現されてしまうが、予測値そのものの分布が知られないかぎり、その範囲をより狭く表現することはきわめて困難なことのようには思えた。

### Synopsis

SUZUKI, KEISUKE: A Study on the Evaluation of Prediction. The Journal of the University of Transportation Economics (*Ryutsu Keizai Ron-syu*) Vol. 8, No. 2, 1973/10 pp. 1-11.

Many methods of evaluation of accuracy of prediction are proposed. The percentage of verification, the coefficient of association,  $\chi^2$ -test, RMS error, a value derived from the theory of entropy I, Theil's inequality coefficient are the examples of them.

In this paper, I proposed tentatively  $U^*$  which is derived from Theil's inequality coefficient, and is defined by

$$U^* = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (\pi_i - \alpha_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n \pi_i^2} + \sqrt{\frac{1}{n} \sum_{i=1}^n \alpha_i^2}} \quad (4.13)$$

where  $\pi_i$  is the  $i$ th prediction  $P_i$  or the logarithm of it,  $\log P_i$ ,  $\alpha_i$  is the  $i$ th actual value  $A_i$  or the logarithm of it,  $\log A_i$ . When the  $\pi_i$  is  $P_i$  and the  $\alpha_i$  is  $A_i$ , the coefficient  $U^*$  is the coefficient of inequality of difference (which is the coefficient of inequality measured by the difference between  $P_i$  and  $A_i$ ) and is equivalent to the Theil's inequality coefficient  $U$ .

And when  $\pi_i$  is  $\log P_i$  and  $\alpha_i$  is  $\log A_i$  the coefficient  $U^*$  become the coefficient of inequality of proportion  $U'$  (which is the coefficient of inequality measured by the proportion of  $P_i$  to  $A_i$ ).

Therefore, the coefficient  $U^*$  generates systematically the coefficient  $U$  (Theil's inequality coefficient), that is the coefficient of inequality of difference, and the coefficient  $U'$ , that is the coefficient of inequality of proportion (When we consider the  $\pi_i$  and  $\alpha_i$  in the formula of  $U^*$  as  $P_i$  and  $A_i$ , respectively,  $U^*$  becomes the  $U$ , and when we consider them as  $\log P_i$  and  $\log A_i$ ,  $U^*$  becomes the  $U'$ ).

If we can have an assumption that

$$P_i = a + bA_i + \varepsilon_i \quad (4.4) \text{ or } (7.1)$$

(where  $a$  and  $b$  are parameters and  $\varepsilon_i$  is the difference between  $P_i$  and  $a + bA_i$ ) in other words, that the condition :

$$P_i \geq 0 \quad (7.2)$$

is accepted (because, if  $\varepsilon_i \geq -(a + bA_i)$ , then we can get the relation expressed by equation (7.2)), and if we consider the difference between  $P_i$  and  $A_i$ , it would be better that the coefficient  $U$  is used.

On the other hand, if we can have an assumption that

that

$$\log P_i = a + b \log A_i + \varepsilon_i \quad (4.10) \text{ or } (7.3)$$

or

$$P_i = (kA^b)e_i \quad (4.11) \text{ or } (7.4)$$

(where  $a$ ,  $b$ , and  $k$  are parameters,  $\varepsilon_i$  is the difference between  $\log p_i$  and  $a + b \log A_i$ , and  $\log e_i$  is  $\varepsilon_i$ ) in other words, that the condition :

$$P_i > 0 \quad (7.5)$$

is accepted (because, even if  $\varepsilon_i \leq -(a + b \log A_i)$ ,  $P_i$  is always positive), and if we consider the proportion of  $P_i$  to  $A_i$  or the value  $\log P_i - \log A_i$ , it would be better to calculate the coefficient  $U'$ .

But, even if we have the condition  $P_i \geq 0$  or  $P_i > 0$ , the both  $U$  and  $U'$  may be used in each case, when we want only to measure the goodness of prediction by a measure of difference or a measure of proportion having no connection with the conditions.

Furthermore, the measures  $s^{(+)*}$  (which may be called "positive standard deviation") and  $s^{(-)*}$  (which may be called "negative standard deviation") defined by equations (5.1) and (5.2) in this paper may be utilized when we want to know the distribution of the difference  $P_i - A_i$  or  $\log P_i - \log A_i$ . And even if the type of the distribution (probability density function) of the predictions are not strictly known, we can quantitatively show or discript the range of the distribution of predictions around the actual values by means of  $s^{(+)*}$  and  $s^{(-)*}$ .

When we express the range of the existence of prediction by  $s^{(+)*}$  and  $s^{(-)*}$ , the range is considerably wide. But, it is thought that it is very difficult to express the range by narrower interval since, in general, the type of the distribution of prediction is not specified.